ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA CCCACGCGTCCGGGCCGGAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCG CTGCTGCCGCCCGCGCGGGGGCCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA CGGCTTGGGAGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATC CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA ACGGCTACTTCAGCTCGCGCAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT GGACGAGGGCCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG GGCTGCACAGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA CGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAACCTGTGTGAGGAAAAACG GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT $\tt GCCCTCCCGCGAAGACCTG{\color{red}{\textbf{TAA}}} TGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC$ CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG AGGAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGAAGC TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA TCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGGTTTGTCCAAACTC ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG ATCTCAATTAGTCAGCAACCCAGTTTT

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTLSKYESSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVLDE
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTOLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343 and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and 313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGGATCCTCTAGAGATCCCTC GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCCCAGCCGTCTAAACGGGAACA GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC<u>ATG</u>GCCCGGAGGAGCGCCTTC CGGGCCGCCGCAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTAC CTGGCAAGCTGCAGGCAGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAAACACCTCAAAATGCTATCT TCTTTAAAACATGTCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG TACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTG GATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACC TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG CAAATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGCACAT GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCCCAGCTCAGGCAGC ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG TGAACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA TGTGTTGAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAG CATGATGGTATAGATTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA TCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT GTCTGGGGGCAGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT GTTTGTTACATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACC AAACAATATAATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT TGGCTTGAAGCAATATAATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT AAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC CGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

CCCCAGCCCACACCTTCACCAGGGCCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC TGCTGTTGCTGCCGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG GAGCTAGCACCGGGTCTGCACCTGCGGGGCCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCCCCTGACTTCTGGGACTTC TGCCTCGGCGTGCCACCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT ${\tt GAACCACAGCGCCTTCTGGGGCATGACCCTGGA}{\tt TGA}{\tt GGGCATTCGCTACCGCCTGGGCACCA}$ TCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT CAATCCATTCTCTGGGACACATGACGCCTGTCCTGTCGCCCCAGAACCTGCTGTCTTGTGAC ACCCACCAGCAGCAGGGCTGCCGCGGTGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG CCGAGGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCC CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC CTTGGGAGGCCAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGCGTCAATGAGTGCGACATCGAG AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGACATGGGTCATCACTGAGGCTG CGGGCACCACGCGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCCAATG CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCGGGCAGGCGAGACTGGCG GAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC ${ t TTGCCCAGGTTGGAGTGCCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA$ AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACCACCTGGC ${ t TAATTTTTGTATTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT$ CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC TAAAACCAAAGTATTGATAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223</pre>

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDC ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPIQGCMHGGRIYPVLGTYWDNCNR CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site. amino acids 398-409

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGAATACCGCCGGCATGGGACCCCACTCAG

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG ${\tt CAGCTTGAACTGCGTGG} \underline{{\tt ATG}} {\tt ACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT}$ GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT GCGCTGCTCCCTGCACTCGGCCTGCTCTGGGGGACCCGGCCAGCTATAGGCTCTGGGGGG CCCCGCTGCAGCCCACACTGGGTGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG CTGCACCCCTGTCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA GGGGACAGGCACTCAGGAGGGCCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT AATAAACACCTGTTGGATAAGCCAAAAAAA

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHPDPPMALSRTPTRQISSSDT DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ

Signal peptide:
amino acids 1-47

N-glycosylation site. amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site. amino acids 22-26, 76-80

N-myristoylation site. amino acids 56-60

Amidation site.
amino acids 70-74

 ${\tt CCCACGCGTCCGAACCTCTCCAGCG} \underline{{\tt ATG}} {\tt GGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGT}$ GCTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACC GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA CGTGCAGGTCACCGGGGGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG CGTGTTCACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGG GCTGGTTCATGGCCTTCACGCGGCAGGGGCGCCCGCCAGGCTTCCCGCAGCCGCCAGAAC CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCACCCGCCGGACCAAGCGCACAC $\mathsf{GGCGGCCCCAGCCCTCACG}$ GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTG $\tt CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT$ GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTA AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCCACCCCCAACTCCCAGCCC CGGAATAAAACCATTTTCCTGC

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLE NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT GAAGTCTTGGCTTATCATTTC@CTGGGGCTCTACTCAGAGGTGTCCAAACTCCTGGCCTGCC CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC AATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC CACCCTCCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTG CCTCTCTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT TCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTC CTGACTGGGATGGCAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATC CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGGCATCGTTCAGGAGCGCATAG TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGGCGCGGTGATATTT GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGAT CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC ACTGCCATACG**TGA**CAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATAATAGTGGGATTTAAAAAAAGTG CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN PWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653 Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655 Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300, 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

ACTTGGAGCAAGCGGCGGCGGCGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCCTCGCCTCCCACGAGCG ATCCCCGAGGAGAGCCGCGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCTGCC ${\tt TCGCTTCCCAGGCGCCGGCGGCTGCAGCCTTGCCCCTCTTGCTCGCCTTGAAA} {\tt ATG}{\tt GAAAAGATGCTCGCAGGCT}$ CTAGGGGCAGACACGCTCGGACCCACCCGCAGACGGCCCTTCTGGAGAGTTCCTGTGAGAACAAGCGGGCAGACC TGGTTTTCATCATGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA TCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACGG GCACCATGACTGGGCTGGCCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCCGGCCCCTGA GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG GGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC AGAAGAAGTTGTGCACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAAA CGTGCACAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCC GTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTG ${\tt CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGAAGCAGTGAAGATTCGTTTGTGT}$ GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACCCACCATGGCTGCGAACACATTTGTGTTA ATAATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT ${ t GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCG$ TGAAGCAGTTTGTCACTGGAATTATAGATTCCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT ${ t ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCCC}$ ${ t ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTTACCCAAG$ ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG CAGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT ${\tt CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT}$ CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT ${ t TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCCC}$ ${\tt TGGAAAATCGCCTGAGATACAGA}{\tt TGA}{\tt AGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT$ GAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAAATCCT TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAATATACTGTGGACAC AACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFIIDSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAVKRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARDTGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHNCAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYALAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGCEHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLINEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCEHSCVSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLAEDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVIDGSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVAHMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDSPAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPLEEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247, 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784, 781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500, 639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464, 540-546, 581-587

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG ${\tt C\underline{ATG}ATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA}$ GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG GCGGTGGAGGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTC AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG ${\tt CATGGGTGCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTA{\color{red}{\bf TGA}{\bf TGA}{\bf CCCCAGAGTCAAGCTGGCTCTCTGGTA{\color{red}{\bf TGA}{\bf TGA}{\bf TGA}{\bf CCCCAGAGTCAAGCTGGCTCTCTGGTA{\color{red}{\bf TGA}{\bf TGA}{\bf TGA}{\bf CCCCAGAGTCAAGCTGGCTCTGTGGTA{\color{red}{\bf TGA}{\bf TGA}{\bf TGA}{\bf TGA}{\bf CCCCAGAGTCAAGCTGGCTGGCTCTGTGGTA{\color{red}{\bf TGA}{\bf TG$ AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGAGA AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCTGACTCCTTATGAAGCCAGCTG CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC CCCCTTGATCTGTACCCCACCCCTATCTAACACCCCCTTGGCTCCCACTCCAGCTCCCTGT ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC TGTTTGTATGAAAAA

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQ WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTL SSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 164-175

CGCCACCACTGCGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTTCCACTTTG TTGAATTGTTCCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA AATACGCAATGGAATTGAAGCCTGCTATTGCÁACATGGGATTTTCAGGAAATGGTGTCACAA TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA TATAATTACATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA AGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGAT AATTCTGAAGAGGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC ACCCACATTATATGAACTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA GGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT GACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTTGCATGGATGTGC ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACTTTATT TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCAT TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTT GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC TTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTTTATCTAGAAAGATTCAAGAAGAAT ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGG<u>TAA</u>ACATAGAGAATG GTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA TGACTCATCAAATTATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTGAAACTAGTCC CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL SNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF WNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSIIAGL LHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA LALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG CTAAGCGAGGCCTCCTCCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA GGTGTGAGTGGGTGTGTGCGGGGGGGGGGGGGTTGATGCAATCCCGATAAGAAATGCTCGGG TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC GGCACTCACAGCCCCGCAGCGCATCCCGGTCGCCGCCCAGCCTCCCGCACCCCCATCGCCGG ${\tt AGCTGCGCCGAGAGCCCCAGGGAGGTGCC}$ ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGCCCCCTCGCCTTCTCGGACGCGGGCC CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC ACGGGCTCTCCAGCTGCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG CGTGCACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA CAGAGGCTTTCTTCCACTCTCATTTCCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATG ${\tt GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAG{\color{red}{TAA}CT}$ GAGACCATGCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCCTGGGGCCCCCATTCTGCTCCCTCGA GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC ${ t TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC$ CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTA AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC TTTTGTATATTAAAATGGAGTTTGTTTGT

MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD MFSSPLETDSMDPFGLVTGLEAVRSPSFEK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site. amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 48-59

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA GAGCAGCCGGCTGCCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG CTGCTGCGCTACCTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAAACTGGAACT CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA TCATAGCAGCCGTAGTAGTGTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT GCTCAGAGGAAAGGCTACTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG $\tt CCGCGGCGGGCGATCACGAGGTCAGGAGTTC\underline{TAG}ACCAGTCTGGCCAATATGGTGAAACCC$ CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC TTGGGAGACAGGAGATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC

MARRSRHRLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE WKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVA LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA TTATATCATTAAGGAAATAGTAACCTTCTCTCTCTCCAATATGCATGACATTTTTTGGACAATG CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTCATCA TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAGCTAAGATGAAGGACATGCCACT CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT TGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT GGACTGGAAAACTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTGCGTGGACC CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA GATATAAATGGCGTAACTCCCAAAGAAGGGGGGTTTATATACTTGTATAGCAACTAACCTAGT GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTTGGTGTCCTGGAAAGCA AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAA TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGGATTATTGGTGTGATATGTCTTATCA AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA AGAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCC**T AA**AAACCACCAAGGAAACCTACTCCAAAAATGAAC

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT FPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE ENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL IKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPD LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESFPSNLNV EAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE YEKNNTTTLMACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLIN LWEAGKEKSTSLKVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583, 608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443, 491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

GCCCGGGACTGCCCAAGGCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG ACCTGATGATATTAGCACTGTGGTA<u>TAG</u>TGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK
ADEPDDISTVV

Signal sequence: amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site. amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 199-203

Casein kinase II phosphorylation site. amino acids 162-166, 175-179

N-myristoylation site. amino acids 37-43, 45-51, 110-116

ACCGAGCCGAGCGGACCGAAGGCGCCCCGAGATGCAGCAAGAGAGATGCTGGCGGGG GGCGTGAGGAGCATGCCCAGCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT ACCGCGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG CCTTCAACAACCTCTCCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC CTACCTGGCCGTCCGCCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGGCCTCAACTACCTGCGCGTGCTCAA TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCGAGTTTGTCCA GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCG CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT GTGTGCCGGGCCGATGGCGACCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT TCCATGCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGGAGAGGCCAACAGCACCCGCGCCACTG TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC CTGGGCGTCGTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG ${\tt ACGCGCCCGCAAGTTCAACATGAAGATGATA}{\tt TGA}{\tt GGCCGGGGGGGGGGGGGGACCCCCG}$ GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCCTTCCCACCTC CTCCCTACCCTTCTACACACGTTCTCTTTTCTCCCTCCCGCCTCCCTGCTGCCCCCCG CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGACGCGACACGCGGCAGAGTCA ATAATTCAATAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFNLRTL GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLA CDCRLLWVFRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

ATCGTAGTCCACCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCCCCAGCC AGGGAGCCGGCCGGAAGCGCGATGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA CGAGGACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTT ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG ACAGGAAGATCCCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC CCCAGCAGTACCTATGGGAGAGGGGGGGGGTGTGCCACCCCTGAAGATGACCCAGGAGAGT GCCCTGATCTTCCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG $\tt GGCAGTCAGGAGGGACGACAAGAAGGAATATTTCATC \underline{\textbf{TAG}} \\ \tt AGGCGCCTGCCCACTTCCTGC$ GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA TTGCCCTCAGCCCTTTCCGTGGCTTCTCTGCATTTGGGTTATTATTATTTTTTGTAACAATCC CAAATCAAATCTGTCTCCAGGCTGGAGAGGCCAGGAGCCCTGGGGTGAGAAAAACA AACAAAAAACA

MGAPAASLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLV TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ KLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDD APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GGGGGTTAGGGAAGGAATCCACCCCCACCCCCAAACCCTTTTCTTCTCCTTTCCTGG CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTCTGTGGCGAGCAGGATGGTCGCTG TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAA<mark>ATG</mark>CTGCTTTGGATTCTGTT GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCTG CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT CAAAATCAGACCCACAGCAGCATAGCGACGGGTAGCTCCAGGAACAACCCTTAGCTAACA GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA TCCTCCCGGGCACTTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG CTGAGGTCCCTGCGTGGACGTGTTCGCTGGGGTCTCGCTCTAAACTCAGCCTGCACAA CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT TAGAAAGGATTTCATGCTCCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGGT GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG TGGCTCTCACTCGCTCTCAGAC<u>TAA</u>GACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG ATACATCCTTCCCCACCGCAGGCACCCCGGGGGGGTGTACCCAAATCCCCGCG CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTCCCCTGTGGATTAG CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGCAGTTGCACGAAGGCATGAATGTAT TGTAAATAAGTAACTTTGACTTCTGAC

MLLWILLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ TFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK IPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKP KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY LDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL VPGLLLVFVTSAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG AHRVYDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 337-348

AGTCGACTGCGTCCCTGTACCCGGCGCCAGCTGTTTCCTGACCCCAGAATAACTCAGGGC TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGTTGGC CGCTGGGCCCGCGGGGGATTCTTGGCAGTTGGGGGGTCCGTCGGGAGCGAGGGCGGAGGGG AAGGGAGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC AGA**ATG**AGGCCGGCGTTCGCCCTGTGCCTCTGGCAGGCGCTCTGGCCCGGGCCGG CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGGGGGCCTGCTACAGCCTGC ACCACGCTACCATGAAGCGGCAGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCCGCTC AGCACCGTGCGTGCGGGCCCGAGCTGCGCGCTGTGCTCCTGCGGGCAGGCCCAGG GCCCGGAGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCCACT GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCCTCCGACCCCGGCGGTCTC GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC GGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGG GCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC ACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGATTCCTCGATGGGGAT CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT CGACTCCTCCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCT TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCA AGGAAGGAGTCTATGGGCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG CAGAGGGTGCCTTGCTGGCGGAGTCCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGA CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC TTGTGTAACTGACAATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAG GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA TGGTGATACTGGGGGACCGGGTAGTGCTGGGGGAGATATTTTCTTATGTTTATTCGGAGAA TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTAT ATTGGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAATTAAAGGATTGTTGAT

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATA TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

CGGACGCGTGGGATTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG GGGGCGTGTGCCGGCGCGCGCGCGCGCGTGCGAAACCCCGAGCGTCTACGCTGCCATGA GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGCTGCCGCCACCCAGCTCTCGCGG CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC CGATATGATTATGTGGCTGTTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA AAACTGCCTACAACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAACCATCACTCGCGATGGGAGTTTG CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTCAGCAGGCGGG CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG TTAACAGTGAACTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC TCTCAGTAGAAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA ATTTTAGAATTGAGTTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGT GTTATTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTFPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295, 305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

 $\tt CGGACGCGTGGGCGGGCGGCCCACGGCGCCCGGGGCTGGGGCGGTCGCTTCTT$ CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCCATGGCCCCCGAAGG GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTCCCAACCTCCCAGGACCTATCTGG CTCCAGCCTCTCCACCTCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAAACCATCCGGGACAACTTTGGAG GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG GTAGAGGTGCTGGAGGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT GAGTGAGGAGCTGGTGGTGGTTTCACAAGCAGCAGGAGCCCCGGACCTCTTCC AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC CTTCCCTGTCCTGGGGGAACAGAGGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAGG GACACGAGGGGCAGCGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC TTTGGCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGG $\tt CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT$ GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG GCCTGCCTAGGCTGCATGGGGGCCAGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC AGAAGACGAGTTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA CGCTGGCTGAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGGCTGTGGCGGCCATG ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA **A**TCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCCTGGGGTAAAAAGTAGC CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGCTGGCAGGCTTCACAATGTGTGA ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTG GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFG IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179, 177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289, 326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA $\tt GCACC{\color{red}{\textbf{ATG}}}{CAGCCCCTGTGGCTCTGCTGGCCACCCCGGG}$ GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAAAGAGGT ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAGAGC TTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACCTGCTGGTGTTCGG CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGCGCTGTCCCCGCGCAGCGCCCGGGCC CTCCAGGCTGGTGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTACAGGTGTCGGTGCAGAGG GAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACCCCTGGACCTTGGGGACTATGGAG CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT CCTGGCTTATGAGTGTGGGCACCTGCCGGCAGCCCCGGAGGCCCTGGCCTTCAAGTGGC ATCAAGGAGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA GTGCAGCTGTGCCTCGGATGGTGCGCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG TAGCCATCGAGGGACTTGACTTGTGTGTTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTAGTGAGCCCTGAATTTGCTT CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCAC TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT AAAA

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG<u>ATG</u>GGGACAAA GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTTGGAAGTTTGACCA AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAGAAAGGGA CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG ACCTCGTCATTCCTGGTGTGAGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC TACACCCCACAGGGCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC ATCCTCCTTCATGCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

CAGCGCGTGGCCGCCGCTGTGGGGACAGC<u>ATG</u>AGCGGCGGTTGGATGGCGCAGGTTGGA GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCT GGAGGCCGCGGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG GCTCGTGCCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC ATGTACCCAGAAAGGGCAATGCCCACCGCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTAGCA GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA CCCAGACTGCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGGTTATTGCAGCTGCTGCGGTGC TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCTTTTGTCCTGGCTCCGAGCCCAGGAGCGC CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC $\tt CTCGCTGCCCTGA GGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA$ GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR TSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218, 224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

CCCACGCGTCCGGTCTCGCTCGCGCAGCGGCGCAGCAGAGGTCGCGCACAGATGCGG CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA TAATTCCATCTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT GCCACCGGGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTCGTCATC CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAG TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG ACGAAGCTGTGAGTGCCGTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG GGCTGCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA GTCTGTATTCACCTCCCAGGTGCCAAGAGAGCCCCCCCCTGCTTCGGACAACCCTGACATA GTTGTTCCTAAGAAAC**TGA**TTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTTGGTTTTAGACAAATGTAAACAA AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAAGNAAAAAA

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT WPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDS VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIHHAHWVLFLRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366, 364-368, 380-384, 467-471, 468-472

N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424, 478-484

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 394-405

CCCACGCGTCCGCTCCCCCCCCCCCCCCCGCTCCGTGCGTCCGTCGGTGGCCTAGAGA TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT AGCGCCCGAGTGTCGGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT GGATCTGTAGAAAAAGAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACCATC TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG GTTATTAGGACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC CTCTTATTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA GCAGGAAAAAAAA

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLV
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226, 299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

ATCCGCAGGTTCCCGCGGACTTGGGGGCCCCCGCTGAGCCCCGGGCGCCCGCAGAAGACTTGT GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCCAGCGAGGGCCTACCACCATGATCACTGGTGT AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA GGCAGAGACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT GGACAACGTGGCTGCCGAGCAGGCACACCACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG CACGGATGATCGAACAGAGGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA ACTCAGGTGATGGAAGTTGGAAATGAAGAG<mark>TAA</mark>CTGATTTATAAAAGCAGGATGTGTTGATT TTAAAATAAAGTGCCTTTATACAATG

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLHILES NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 218-222

Casein kinase II phosphorylation site. amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site. amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature. amino acids 50-65

 $\tt CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT$ TAAATTTCAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT CCACCGCTGTGAAGGCCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGGCTGTGGAGCTGCCAGC GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAGAGCCAAAAGGTCCTCATCCA ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA GTCCCAGAGGGTGTCAGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA GGATTGCCCTTCTGGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG GAGGTGCTGCACAAGGGCGTATGGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG TGTGGCTGTCATCTGCTCAGTGTAGGTGGCCATCATCTAATCTGTTGAGTGCCTGAATAGAA GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCCTTCAGAGTTGG ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG TATGTGTTCAAA

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238, 267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143, 180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC CCACGCGTCCGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG GTGAAGGCCATCGACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG CGTCTTCGGCCTCTTCCGGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACAAGCCTTACTTGGTGACCTTCGACC TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

In the first of the first that the start of the first the first that the first th

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

CCCACGCGTCCGCTGGTGTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA AAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC CTGCTTCTCCCGTTACTGATCGTCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCCTAA GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTCATACCTTTGT GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG TTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG TCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA CCTGGCAGTTTTAAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA AAGCGCAA<u>TAA</u>GCACCTAGTTTTCTGAAAACTGATTTACCAGGTTTAGGTTGATGTCATCTA ATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTTTCTAATTATCCCCATTTCTTCAATA TCATTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT TGAAGGACTATATCTAGTGGTATTTCACAATGAATATCATGAACTCTCAATGGGTAGGTTTC ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA

MKFLLDILLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

AGGGAGGAGCACCGACTGCGCCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA CCTTGCGTGACAGAGTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC AGGAAGCAGAACTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAGTACTTGCGAGA AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGA<u>TAA</u>ACTACCTTCCCAAAAGAGAACAT CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC TTTTGTTTTTGGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA AAAATATTATATATAAAAGTAAAAAAAAA

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH GPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT EIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWKIFKSDSEVAGYIRQ AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352, 353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT TTTTCCCTTTCCTAACAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGAGAAAGTATGTTAAAAATA ${\tt GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGG} \underline{{\tt ATG}} {\tt GGGACCCTGGGTC}$ AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTT CCTCACCGCCCCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT CCACCAAGGGACGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG CCCCTCATCGTGCAGCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCT CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA ${\tt TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA}$ GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT TTTGTCTCCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG CAGTATCACCACCCGCCCGATGACTCTGCCCTGTGTGCCCTTCCCTATCCGGGCCATCAACTT GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC ${ t ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG$ ${ t GCAACTTTATTTTCTTGGGGAACAAAGG}{ t TGA}$ AATGGGGAGGTAAGAAGGGGTTAATTTTGTG ACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA TTTCAATATTTCCCAAACTTTAAGAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSL
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387, 384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

 ${\tt AGGCTCCCGCGCGCGGGTGAGTGCGGACTGGAGTGGGAACCCCGGGTCCCCGCGCTTAGAGAACACGCGATGACCA}$ GCAGGCTGGACTGGAGCACCCTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCCAAGGGCTGGA GGAGGGACCGCCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCCTCACCACCTATGTTCCGTGGAACCTGCATG AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCCTGATGGCCGCAGAGATCG GGCTGTGGGTGATTCTGCGTCCAGGCCCCTACATCTGCAGTGAGATGGACCTCGGGGGGCTTGCCCAGCTGGCTAC TCCAAGACCCTGGCATGAGGCTGAGGACAACTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC GTTCCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACTGC TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGATGGAGT ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCT TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGGAGAATATTGATGACCAGCGCAAAG GCTTAATTGGAAATCTCTATCTGAATGATTCACCCCTGAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA GCTTCTTTCAGAGGTTCGGCCTGGACAAATGGNGTTCCCTCCCAGAAACACCCCACATTACCTGCTTTCTTCGG GTAGCTTGTCCATCAGCTCCACGCCTTGTGACACCTTTCTGAAGCTGGAGGGGCTGGGAGAAGGGGGTTGTATTCA TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA GCAGCGGAATCAACCAGGTCATCGTTTTTGAGGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC ${ t ACCTGGGCAGGAACCAGTACATTAAG{ t TGA}{ t GCGGTGGCACCCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC}$ ${ t ACTGGGGGCTACAGTCTGCCCCTGTCTCAGCTCAAAACCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCCC}$ TGGCTTTGTTGATGATGGCTTTCCTACAGCCCTGCTCTTGTGCCGAGGCTGTCGGGCTGTCTCTAGGGTGGGAGC TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT ${ t TTTATCCCCGAAATCCTGGGTGTGTCACCAGTGTAGAGGGGTGGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT$ TCCCTTCCCACTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCCTGCGTCTTCC CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCA CATCCAGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCC AGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA GGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG ACAGAAGGCCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCCGAACAGCAGGGGGCAGAGCAGCCCTCCTTC GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTTCTGGGGCCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA GTTGCAGTAAAGCTATAACCTTGAATCACAA

MTTWSLRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLEAFVLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ
YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDGLSKGIVQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDSSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE
TSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQRFGLDKWXSLPETPTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNOYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315, 320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC $\tt CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCC{\color{red} \underline{ATG}} GCTCCC$ AAGAAGCTGTCCTGCCTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGCCC CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCGAGCCCCAAGATGATGCTTGGACCTGTG ${ t ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT}$ TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG AGACAAACTATTTTTGACGGGGAAACTGGGTCCAAACTGGATATCTTGGTGGAGAACATGG GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG ACCAAGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTCCTAGGGGAGCCCTCAACAAATTA CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG ${\tt TGCCTCTGAACCAATGGAGTTAAGTGGGCAC}{\tt TGA}{\tt AAGGTAGGCCGGGCATGGTGGCTCATGC}$ CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG ${ t ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGGCTGAGGGCAGGAGAATTGCTTGAATCC}$ AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA GACACTCCATCTCAAAAAAAAAAAA

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554, 603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233, 231-237, 274-280, 296-300, 307-313, 447-453, 484-490

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT. TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT GCGTTCCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC CTTGGACCGCCTGCCAGCCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA AATATTCCCTTTGCAAATGGGATT**TAA**ACTAAGATAATATATGCACAGTGATGTGCAGGAAC AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG ATACATCTTTTAAAATAAAACAGAGAGGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTTAAATTGTTTTGTAACTTGGAT GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT ТТСТТАСТААААААААААААААА

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLFMLSG
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLESLPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTTACACACATAAGGAT GCTCAGTGCTTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACACAAC GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC ${\tt CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACTGACGTTTCCCTGGAGGTGTCCAGAAA}$ ${\tt GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAA} \underline{{\tt ATG}} {\tt GAGCTTGTA}$ AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTTCCTATGTCCAGCTT GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG CAAGTCTGCAGTAAAAACGACTATGTTCCTGTATTTGAATCATCAGCAGTACATTGACGTTTCAAATAGTTACT GGCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCCTAGAAATAGAC AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT TACCGGGGATTTTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA GACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTGAATCCAATTCA TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTTGTTCAAGTTAGTCTGCAC ACCTCAGATCCAAATTTGGTGGTGTTTCTTGATACCTGTAGAGCCTCTCCCACCTCTGACTTTGCATCTCCAACC TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGA TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG AAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTCAGCAT GAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTG AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG CAGAACTAT<u>TAA</u>CTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACGGCCTGCATGTAAAAAAA

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI ERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSTLT FQIVTDSARIQRTVFVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374, 394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383, 408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGC TGCTGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTTGGACTGTACTATTCCCTTTTTGA ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA TACCTGGCGATCCCAGATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTTCCCTATTTCCCACTA AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCCCAGTGCACTTTGCCATTAAAGTC TCTTCACATTGATTTGTTTCCATGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT TTTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAATGTGAAGGCCTAGGACATTA GTTTTTCTTCAATTAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA **ATGTAA**

MRPQELPRLAFPLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTTG CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA AACTGCAACACAAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG AGAAAAAATTGATTTGTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA TTACCCTGCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAAAGAAGAAGTCCACCCTT GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA ATCACTGTTTAGAACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTATCTGA GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTCA GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTTGCTGAGACTAATCTT ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA GCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA AGCATTTAGAAAACTT

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR LLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQF AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF GAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

AGATG GCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGGCTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC TGATGACGTGCAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA TTGGTCTAATGACTGCCAATCATTTGCCCCTATCTATGCTGACCTCTCCCTTAAATACAACT GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC AAAGTGAGCACATCACCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA GGAGGCAATGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTCAGA ${\tt TGGGGAAAACAAGAAGGATAAA}$ ${\tt TAA}$ ${\tt GATCCTCACTTTGGCAGTGCTTCCTCTCTGTCAATT}$ CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG TCATAGGACGGAGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG TTTTTATCATTGAAGACAATATTGAACAACCCCCTATTTTGTGGGGGATTGAGAAGGGGTGAA TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT TCCTTTTG

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCDFD WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGGACGTTTGCCCTG $\tt GGGCCCCAGCCTGGCCCCGGGTCACCCTGGCATGAGGAG{\color{red} ATG} GGCCTGTTGCTCCTGGTCCC{\color{blue} A}$ TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGGGGCCTGGACTGGTGCAACGCGGG CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT GACGCTGACAGAGGCAAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC CGTTTTCCCTTGTGGGGTTGGAGCCATTTTAACTGTTTTATACTTCTCAATTTAAATTTTCT TTAAACATTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTTGTGGGCCTTCTG GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGCCCAGAGTGGGC GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCCTTTGGGCCTCTCCTGCC CACCTCTACTTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG GGAGGGCCCTCAGGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR HRRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site. amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

GGAGAGCGGAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC CAGGGGAGGTGCACCAGGCGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA ${\tt CTTCCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA}\ .$ CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC GGCGACGGCTGGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA CATACGGGACTCGGTGAGCGCCTGGGACACGTACGACACGGACCGCGACGGCGTGTGG GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGAC GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGCGTTTCCGGGTGGC CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGG AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAA GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA GGAGCCGGCGTGGGTGCAGACGGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCCCCTGCCCAGGACCAGCCC CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG ${\tt ACCTGACCCGGCACCACGATGAGCTG} {\color{blue}{\textbf{TGA}}} {\color{blue}{\textbf{GCCCGCGCCACCTGCCACAGCCTCAGAGGCCCG}}$ GCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTC TGTCCCTGTCACACCCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA GAACCGCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT GGCCCCAGCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA AAAAAAAAAAAA

MMWRPSVLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ FRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG GCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCAG CCTGTCTGTCGTCGTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTG GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG CGGGTCGTCTGTCTCTCTCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCGGGGCTCT GAGGAGGTGACGCGCGGGGCCTCCCGCACCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG GTGTGAGCAGCCTATCAGTCACC<u>ATG</u>TCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG TGTCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC TGGTCGAGAAAACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGG GCCGATTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAA AAACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT CGAGGAAGCAGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTC ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT GATGCCATTTCCTTCACTGTTAGAAATGTGTTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA CTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGGCTTGGGCACCTCTGGATGACCTG AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTCACAGGATT ${\tt AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGTTTCTTAGAATCCCAGCAAT}$ **AA**TGGTAACATTTTGACAACTGAAAGAAAAGTACAAGGGGGATCCAGTGTGTAAATTGTATT CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACTATTAAGTATGTCAAC AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424, 425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211, 239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

CGCCGCGCTCCCGCACCCGCGCCCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC GGCGGCCTCCCGGCGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG GGGCGGCGGCGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC TGCTGGCGGCGGCGCCCACGGCCCCCGCGCCCCGACGGCGACCTCGGCTCCAGTC AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA GGTTGAGGAACTGATGGAGGACACGCAGCAAATTGCGCAGCGCGGTGGAAGAGATGGAGG CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTAT TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG TGGGAGACGAAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCACAGCCACAGCCTG GTGTATGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGGAGCCTGCGGCTGCCGCCGCT GCACTGCTGGGAGGGGAAGAGATT**TAG**ATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA TCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGC TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG AGTCTCCCTCTGATTGGTTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTTGCAAACATCAA CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCC AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG GCTCAGAGACTGCAAGCTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT CTGGTTGTGACTCTAAGCTCAGTGCTCTCCCACTACCCCACACCAGCCTTGGTGCCACCAA AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG GTCAAACTAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT TGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC ·TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA TAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

CACACATACACCTTCCTCTCCTTCACTGAAGACTCACAGTCACTCTGTGAGCAGGTCATAGAAAAGGACAC TAAAGCCTTAAGGACAGGCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAACATGGGAGGG CCAGGCACGGTGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAGGAG TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC AGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT GCAATGACTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAGAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCTTGGGCCCTGGCAGGGTTGCTGACCCTTACCCTGCAAAACACA ${\tt AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACC{\tt ATG}}{\tt AGGCTTCTCGTGGCCCC}$ ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCCTCA GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCCGCTCGTCCTACCGCGAGGCTACCACTGTGGACTGCAATGA CCTATTCCTGACGGCAGTCCCCCGGGCACTCCCCGCAGGCACAGACCCTGCTCCTGCAGAGCAACAGCATTGT CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCCAGAACAGCTTTTCGGA TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT GGAGGACCACAGCTTTGCAGGGCTGGCCAGCCTACAGGAACTCTATCTCAACCACAACCAGCTCTACCGCATCGC CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCCAACCTCCAACCTCCTGAGGGCCATTGACAG CCGCTGGTTTGAAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT GAACTTCCGGCCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCCAGGCGGGCACT GGAACAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGC CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT GAACCTCCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCCTTCATCCACCCCCGCGCCTTCCA CCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACACGCTCTCAGTGCCTTGCACCAGCAGACGGTGGAGTC CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTCATCCGCTGGGCCAATGC GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCTCATCTCCCCACGAAGCTTCCCCCCAAG CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACCCGAGATCTACTG GGAGCTGCGGAGGGTGACAGCAGAAGAGGCCAGGGCTATACACCTGTGTGGCCCCAGAACCTGGTGGGGGCTGACAC CAACATTACCCGCCTCCTTCAGGCCACGGAGTACTGGGCCTGCCAGGTGGCCTTTGCTGATGCCCACACCCA GTTGGCTTGTGTATGGGCCAGGACCAAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCCTGGGCTCAT $\tt TGCCATCCTGGCTGTCCTTCTCCTGGCAGCTGGGCTAGCGGCCCACCTTGGCACAGGCCAACCCAGGAA$ GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTTCTGGGGGCTGGAGTGCCCCTTCTGTCCGGGTTGT GTCTGCTCCCTCGTCCTGGCAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC ACCATTGTCTCAAAATTCT<u>TGA</u>AGCTCAGCCTGTTCTCAGCAGTAGAAATCACTAGGACTACTTTTTACCAAA AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGC CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCATCTT $\tt CTCTCTGCCCAGAGGCTCCTGGGCTTGGCTTGCCCTACCTGTGTCCCCGGGCTGCACCCCTTCCTCTC$ TCTTTCTCTGTACAGTCTCAGTTGCTTGCTCTTGTGCCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC CTCGGGGGGCTGCCTCAATGTGGGAGTGACCCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCCAGGAA CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA AAAA

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ LARVPRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLP ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTDHCLPLISPRSFPPSLQ VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAEEAGLYT CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVV SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146, 243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGGACATTGTGTACCGCCT CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCACCC CCTGGCCACACTCTTCAAGATCCTGGCGTCCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA GCGCTTCGCCGTCTTCCTGTCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTCATGCTCAGTGGCAT CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCCTCAAGCTGGAGCTGATCCCGACGTGACCATCCCGCC CAGCATTGCCCAGCTCACGGGCCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCCTCAACAGCCTCAA GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCCACTCCATCTTCAGCCT ${\tt CCACAACCTGCAGGAGATTGACCTCAAGGACAACCATCGAGGAGATCATCAGCTTCCAGCACCT}$ GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA CTACCTGGACCTCAGCCAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG CAACCGGCTGGAGTGCCTGTGGAGCTGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGGCTGACAAGGAGCAGGCCTGAGCGAG GCCGGCCCAGCACAGCAGCAGCAGGACCGCTGCCCAGTCCTCAGGCCCGGAGGGGCAGGCCTAGCTTCTCCCAG AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCAGGAGCCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGA GGACAGTATCTGTGGGGCTGGCCCCTTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG AGATAACTTATACATTCCCAAGAAAGTTCAGCCCAGATGGAAGGTGTTCAGGGAAAGGTGGGCTGCCTTTTCCCC CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCACGGGAGAGCAGGCCTCCAGCTGGA AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGACGT TGAACTGTGTTTCCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGGATCTGGTGTGACCTTGGTCCAGGAGTT ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTCGGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGAGGAG ${\tt CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTCCCCACCTCCTGCGGCATGGGTGTGT}.$ CCAGTGCCACCGCTGGCTCCGCTGCTTCCATCAGCCCTGTCGCCACCTGGTCCTTCATGAAGAGCAGACACTTA GAGGCTGGTCGGGAATGGGGAGGTCGCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGC CTGGAGTGCACACAGCCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT AGAAGGGTCCCCGCCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT ${\tt CCATCCGTCTGTCCATTTGTGTTTTCTGCGTCGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAG$ CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTCACAGTATCAAATAAA ATCTATAACAGAAAAAAAAAAAAAA

MRQTIIKVIKFILIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIISFQ HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLT FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368, 398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTC<u>ATG</u>GAGCTGGCACTGCGGCGCTCTCCCGT CCCGCGGTGGTTGCTGCTGCCGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT GGCCCACAGAGGAGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGC CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT GTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC ATTGGTCTAGAGCTTTATAAGGCCATTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT TGCCTTGGGTGATTCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGT ACAGCATGTCTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT GATCATTGAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACTAAAAGCACTCCCA CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT GATGAGACTGGTGACTCAGCAAGAA**TAG**GATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAAGATTTTTTAAAAAATTGAT TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNKGLYRE ATELWGKAEMIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDALS QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDLIVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP SDQGDMALKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175, 187-193, 195-201, 331-337, 332-338, 360-366

GGCCGCGGGAGAGGCC<u>ATG</u>GGCGCGCGCGGGGCGCTGCTGCTGCTGCTGCTGCTC GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCGTTATCAGGACCATGCGGCCGA CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG CACTCACGGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG AGAGGATGAGGCACTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT ${\tt TCTCTGGGCTCTCCCACTCCTGGGGCCGGTC}{\tt TGA}{\tt GCCTACCTGAGCCCATGCAGCCTGGGGC}{\tt CTGGGGCC}{\tt CTGGGGGCC}{\tt CTGGG$ CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTTGGTAATAAACACATTCCAGTTGA

MGARGALLLALLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245, 259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 252-263,

Serine proteases, trypsin family, histidine active site. amino acids 78-84

CTTTGCCCTCATCCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAG CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC ${ t ATCCCTACCAGCCTTCCACAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGGACTGCACCGT}$ TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT GCATCTGGGGGTAACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGAC TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGT ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG GGACAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGT GGGGCCGGGTGTTGGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG CCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA TTGTTGACTATATCAGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAG GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTCAGCAACA GAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTTGGGGGGATCCTA TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTTGGCTTTCTCAACCCAAG GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG GAAGCCCTGCTGAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA TGCTGTGAGCTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA TTTCACTTGATATTCATTCCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTTACTCT TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG ACTACTCTTGTCTTCCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA TGTAAAAAA

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225, 248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488, 521-527, 533-539, 549-555

 ${\tt GGGCTGCTCGGCGGAACAGTGCTCGGC} \underline{{\tt ATG}} {\tt GCAGGGATTCCAGGGCTCCTCTTCTC}$ TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTG GCCTGCATACCGCCTCCTGTCGTCTTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACT TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACT CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG CAGCCGCACAGAGACGCAGGTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC AGGTTCAGCATTTTTGGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTT ATCCACGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCCACTGCA TACACGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA ATTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCCACAAGAGAAAA TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTC TGGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA CCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTTTCAGG GCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGAATCACTC ${f TGA}$ CACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTTTAGGAGAGGCC CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG GTTTGTGTATCATATCATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG CAAACTTTGATTTTATTTCATCTGAACTTGTTTCAAAGATTTATATTAAATATTTGGCATA CAAGAGATATGAAAAAAAAAAAAAAA

MAGIPGLLFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

 ${\tt GCATCGCCCTGGGTCTCTCGAGCCTGCTGCTCCCCCGCCCCACCAGCC}$ CTGGAGCGCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTG GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAACCATACCTGTTCTCTGTGCTGCTGGG CCCACCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC CCCACCTCAGACCTGCAGAAGCTGAAGGTTCCTATCATCGACTCGGAAGTCTGCAGCCAT CTGTACTGGCGGGGAGCAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT GGAGGGGAGCGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG GCGCCTGCTGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA GCTCCGCGGGCGCTCAGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCAC GACTTCCGGCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT CTCCAAAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG SGAAARS

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA GATGGGACGAGGGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC AACAAGGAGCGCGGCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT GCCGCTGGCCATGGAGGGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGTGGAAACCCAGGCC CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCCTTGCCAGGTGCAGAGGGCCCTGACA AGCCTAGCGTTGTCTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCCTCTCCTGGGA. $\tt CTACTGCTCCTCCTCTGGTGTTGGCTGGAATCTTC\underline{TGA}{A}TGGGATACCACTCAAAGGG$ TGAAGAGGTCAGCTGTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATA CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGGCTCAGGAGGCCGCCTGAGGACTGCACACC CACTGCCTACCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG TGTAGCTGGGGATGCGGGATTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC TTTGAGTGGGGGAGGCAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT GTCCAACCTGTGAAA

MHGSCSFLMLLLPLLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEE LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEV SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLAS VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV SGLNSGPGHVWGPLLGLLLLPPLVLAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

GTAACTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTCGGTCATTCTCCAAGTTATGGTGGACGT ${\tt ACTTCTGTTGTTCTCCCTCTGCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTCAT}$ CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAATGAATTGGAGACCATTCC AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA ACATCTGAAAGAGTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATT TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG GCAGGAACTTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATAC ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT GGATCTGAAGAACAATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAG GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATT TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGA TGGCTTTGTGTGTGATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTC CAATTTGAGTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCCAATGACTTTTGCTTGGAAAAAAGACAATGA ACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCAT CCGAGCTGGGGCCATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGA GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGC AACTCTGACTGTCCTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC CGTCCTACAGTGCATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGT AACCGAGAGGCACTTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA ATACACATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC CTGCGACTCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGC CGTGGTTTGCTGTGGGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACACGCGGAGGAATGAAGA TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTT AGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGG ${\tt ATTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC}$ CACAGATCTGTTCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA TCCTTTTGAAACATATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTA CATAAAGAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC GTGGCCTTCACATGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTG TCTAAACAAGTCCTCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGG TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCC AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTC GAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTTAAACTA TTTTTTAACTTTGTTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT TTATAATGCCAGATTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP GYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKM QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE FCOKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFOSFVNASCAHPOLLKGRSIFAVSPDGFVCD DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA MARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAON SAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPOMTAPSLDDDG WATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLAD RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP MYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDA YSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQS YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAAGAAGAACAAC GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA CCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTTTTCTC TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCTGGATGTTGCT ${\tt GTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA{\tt ATG}{\tt TTGAACAAGAT}$ GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCTGCTTGTGGTGCT GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA ${\tt CCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT}$ GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGGTATTTGAACCTTGC CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATCTTTCTGGGAATCA TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGAT TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG ${\tt TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG}$ TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAAATGTCGGGCCTCCACATC CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA ${\tt TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC}$ AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCAGATAACAATGTGGGTCCCACTCC AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCGACAGAGAAAACCTT ${\tt CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGGTCATGAAGACTACCAAAATCATCAT}$ ${\tt TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA}$ TCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC CATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT CAACCACAACAACAGTTAACACAATAAATTCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAAACAA AAAAGAAAAGAAATTTATTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

 ${\tt AGCCGACGCTGCTCAAGCTGCTAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGGCA}$ GAGGGCGGGCGTGCACCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCCGCCGGCTGGGAGCTTCGGGTAGA GCGGTGCTGGGGGGCGCCGGTCCGACAGCGGCGGCGGCGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC GAGCGCCCATGCCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT $\tt CCCGAGCCACTCCCGTCGTCGGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC$ ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA GTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG TTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCCCTACAG CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA $\hbox{\tt CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT}$ ${ t CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA}$ ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG AACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC AATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTC CAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCATTTAAATACA ${ t AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGT$ GATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC ${ t ATCTGCTCAGCTGCCAGCAGTGATTCCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTACTGCATGAT$ GCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG ${\tt CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTC.}$ AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC ATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC ${ t TTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA$ GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGCAACTCTGACTGTC ATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCAC TTTTTTGCAGCAGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG ATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCT ${\tt CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGT}$ GTGGTGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAGGCGGAGGAATGAAGATTGCAGCATTACC ${\tt AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAG.}$ GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGATTTTTCTTACCA CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTC ${\tt CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACA}$ TATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC TCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTAT ${ t TTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTCAGAGGAAGAAGGAAAGGACAGATTTTCAGGAA}$ GAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC $\overline{\text{CTTATGTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAATGCTTTATTATACAGATGAACCAAAATTAC}$ AAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTAACTTTG TTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTATAATGCCAGA TTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY FDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK KLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA RLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA TVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQ DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS SFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT Signal sequence:

amina agida 1.27

amino acids 1-27
Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

CAAAACTTGCGTCGCGGAGAGCCCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC TGGGGGAGCGCGTTCGGCCTGTGGGCCCCCTCGGCGCGGGGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCCTGGGTGGTCCCCTATCCCTCCTTTATATA GAAACCTTCCACACTGGGAAGGCAGCGGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG GCGCACAGCATTCCGAGTTTACAGATTTTTACAGATACCAA<u>ATG</u>GAAGGCGAGGAGGCAGAACAGCCTGCCTGGT TCCATCAGCCCTGGCGCCCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCCAGAGCCGGGTGCTGC TGCTCCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA GTGGCGGCCACAGCCTGAGCCCCGAAGAAACGAATTTGCGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG AGCCCGGGCCTGGCCCAGCCGCGTCAGCTGCCCCCGAGACTGTGCCTGTTCCCAGGAGGGCGTCGTGGACTGTG CTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC TGACCTTGGCACCCCGCTTCCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC TGCCGGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCCTCATCCTGTCCAGCAACTTCCTGCGCCACGTGC CCAAGCACCTGCCGCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCAGCTGGGCTGC CGCGCAGCCTGGTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCC GCACCCTCATGATCCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCCTGGAGG AGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGGTGCACCGCGACGCCTTCCGCAAGCTGCGCCTGCTGC GCTCGCTGGACCTGTCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG TCAAGCGCAATGAGCTGGCTGGCCTTGGCACGAGGGGCGCTGGCGGCATGGCTCAGCTGCGTGAGCTGTACCTCA ${\tt CCAGCAACCGACTGCGAGCCGAGCCCTGGGCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA}$ TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAACA AGATTAGTGCGGTGCCCGCCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAACAAGC TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT $AGGAAGAGGAAACAAGA{\color{red}{TAG}}$ TGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGGACTCTTTTCTGC AGCACACGCCTGTGTGCTGTGAGCCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC GGAACTCACAAAAGCTGGCTTTTATTCCTTTCCCATCCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC CAGGCACTTTTCCAATGGGCAAGCCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEE

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

CAGTAGCCCGGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCACCACCC CAACCTGTTCCTCGCGCGCCACTGCGCTGCGCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGT TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTCGCCAGTCTTGGGGACAGTGTCAGCC ${\tt TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTCATCCTGGTTATGCTGG}$ AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG $\tt CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT$ GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGG CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA TAATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCAACCACCGGACTGACAACTATAGC TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTGCCTGTCATTCAGGCA CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCCACGGAGCAGCCCTGTG GGGAAGAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA AAGA<u>TGA</u>TTAAAGGGTTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG AGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGGTGGCGACCAGCTG TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG ${\tt CATTTCCCTCTTGACCTCCTAATGGAGAGGGGATTGAAAGGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA}\\$ AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC ATTTTCATCGGGTGCATTCTCTCTGCTTCGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC ACACCGGCAGACCTTTCCTTCACCTCATCAGTATGATTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTGGTAAAGGCAGGGCTGG ${f AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTCATT$ TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTAATGTTTT

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDNRVQTDPQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242, 421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGA CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTAT GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT CCTGGATTAGGAGGACTGACAATACTACATATCATTAAGCATGGGCCTCGCTTAGAAGTT GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGCTTACAACAACATAAC CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT ATATCCATGTGTCTTCTCTTCCTATGGAAAAATATCAACCCTACAAAGTTATAAAACAGAA ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAGTGAACTTTCATGG GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA TGTGGAAACTTTACATTGTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTTATTAGTG TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACTATGCCTTCTCTTTTTTTCAATCACC AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA

MWLKVFTTFLSFATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GCAAGCGGCGAA<u>ATG</u>GCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTT GCTTTGGGGTGCTCCCTGGACGCACGGCGGGGGGGGGACGACGTTCGCGTCATCACGGACGAGA ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC TTCCTACTATTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC AGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCTTCAAAAAAAT TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG TTTCACTGTGTGAAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTGTCTCAAG AAATGTGTATTTCAGTGACAATTTCGTGGTCTTTTTTAGAGGTATATTCCAAAATTTCCTTGT TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC GTTTCAAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATT TTGGCCACTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG CATCTTCTTGTATATGTCTTAAATGTATTTTTGTCCTCATATACAGAAAGTTCTTAATTGAT TTTACAGTCTGTAATGCTTGATGTTTTAAAATAATAACATTTTTTATATTTTTTTAAAAGACAA ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAA GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT CACCTTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGAATCCATTAGATTTACAG TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCATT AAAGAGTTTGGATGTGTAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAACCT

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSE EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG ${\tt AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACC} \underline{{\tt ATG}} \underline{{\tt AAACTCTGGGTGTCTGC}}$ ATTGCTGATGGCCTGGTTTTGGTGTCCTGAGCTGTGCAGGCCGAATTCTTCACCTCTATTG GGCACATGACTGACTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCCTGTGAATGCCTACAAACTGG TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCCTGCAGGACTCAGCTGCA GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCCACTGATGAGGACGAGATAGG AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG AAAGCAGCTTGATGCCGGGGAGGGCCACCACCAACCAAGTCACAGGTGCTGGACTACCTCA GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA GGAAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCA TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCGTGGG GAGGGTGTCAAACTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGGGGGCGAGTGGGACAGCCCGCACA TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATCCTGTTGTGGCCCGAGTAAATC GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCCC TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCCTTCCCCCTTCCTGGTC CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTATGTTCCTATCAGGCT GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT GTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTTTTAT

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYHTV LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLR YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLF CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTG VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270, 346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

AGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGAAGCCAGCGGGCCTTGCTCTTTTGGCCTCATTGACCC CAGGTTCTCTGGTTAAAACTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC CCTGGGGCACCCACCTGGCAGGGCCTACCACC<u>ATG</u>CGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC CCCTCATCTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCCTGCGGGTTTCCTGGATCCAGGGGGAGAGAAG ATCCCTGTGTCGAGGCTGTAGGGGAGCGAGGGGCCACAGAATCCAGATTCGAGAGCTCGGCTAGACCAAAGTG ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCCTGACCTCCCGAGCTACACTGTCCA CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTTACTCTACTTCACTGGGCAGCGGGGG GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC $\tt CCCGCCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTG$ GCGCAGGCGAGCCCGGTACTGTCATGGGGGCTTTGGCTACCTGTTGTCACGGAGTCTCCTGCTTCGTCTGC GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCCTGACGAGTGGCTTGGACGCTGCCTCATTG ACTCTCTGGGCGTCGGCTGTCTCACAGCACCAGGGGCAGCAGTATCGCTCATTTGAACTGGCCAAAATAGGG ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTTCGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGT ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA TCCGGAACCTGACCGTGCTGACCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCTTCCA CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG CTCCCAAGTGCCCACTACAGGGGGGTAGCAGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA ${\tt ATCGGCGCTATCAGCCCCGCCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC}$ GGGGCATGGAGTACACCCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGCCCTGGCTCGCA AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCCTCGAGGCGTTTGCAGCCAATGTCC TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTCG $\tt CTGTGCGAGCAGGGCCCCTTCCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCTGTGGACACTCTCT$ TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCCTCAACCGCTGTCGCATGAATGCCATCTCTGGCT GGCAGGCCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCACCACAGAGATCACCCCCAGGGCCCC CGGGGGCTGGCCCTGACCCCCCCTCCCTGGTGCTGACCCCTCCGGGGGGGCTCCTATAGGGGGGAGATTTG ${\tt ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAAC}$ ACCTCTTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCCTAGGGGGCCCGTGCCCAGCTGGCTATGGCTCTCTTTG CCAGGAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTAAATATGAAAATGTTATTAA ACATGTCTTCTGCC

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD QSDEDFKPRIVPYYRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAVNRTV AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV LGWDYFTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLLNGYR RFDPARGMEYTLDLLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLPLL VAEAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAAELERRYPG TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP VHFQEFNPALSPQRSPPGPPGAGPDPPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA RARLAGELAGQEEEEALEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR CRLSNLEGLGGRAQLAMALFEQEQANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC $\tt GTTCTAGACGCGGGAAAA{\color{red} \underline{ATG}} CTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG$ AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA TGCACCACCATGAGCATCACCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACC CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTG GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTTGTTAA AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTTTGGGCATATTTTCAAT GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA TACAGCTGTATGTTTCTTTTTCTTAATTTGGTGGCACTGGTATAACCACACATTAAAG GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT TGTGATTAAAGTAAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAA ATTAAAGTGAAAGTTGAAAAAT

MLSESSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK AYKYAFDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 14-25

Cell attachment sequence.

amino acids 247-250

CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTCTCCCAAATGTTCTTATGGACTGTTGCT GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC AGCTGCTACTTCTTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA TGGGTGGACGCCACACTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG AACAAAGGAAAATCTCTT<u>TAA</u>GAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA AGAACATGGCCACACCCACCCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC AAAAA

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

GGGACTACAAGCCGCGCGCGCTGCCGCTGGCCCCTCAGCAACCCTCGAC<u>ATG</u>GCGCTGAGGCGGCCACCGCGAC TCCGGCTCTGCGCTGCCTGACTTCTTCCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC TCAAATCCAGCAATCGAACCCCAGTGGTACAGGAATTTGAAAGTGTGGAACTGTCTTGCATCATTACGGATTCGC AGACAAGTGACCCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAACAAAA TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC ACTGCCAGGAGTGAGGGCCACCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCCACGGATT CCAGAGCCAATCCCAGATTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG TCACGTTGGGCATCTGCTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA ACCCAGGGAAACCAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTCAGAAGCTTTTCGTTTTTGGCCAAAGTTGACCA CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCCGCATGAGTATTAGG GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCGTGCTGGGCCCCTGTGAAGCCAGCATGTTCACCACTGGTCGTT CAGCAGCCACGACACCATGTGAGATGGCGAGGTGGCTGGACAGCAGCAGCAGCGCATCCCGGCGGGAACCCA GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCCTGCTTGTCCAACAGGGTGTCAGGATTTAAGGAAA TACATCTAAATTTTTGCTAAGGATGTATTTTGATTATTGAAAAGAAAATTTCTATTTAAACTGTAAATATATTGT TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT ${\tt CACACAAGTTTTAGCCTTTTTCACAAGGGAACTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT}$ TAAAAATTCCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT ${\tt CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCCGCTCTAGCTCACTGTTGCCTGCTGTCTGCCAGGAGGCCCT}$ GCCATCCTTGGGCCCTGGCAGTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTAAAAGAAAATGGAT GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG TGAAACGCCTGAATCAAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCCGCAGGCCGCCTGGCAGAGGCA GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGCCATCCTGGATGCTTAGCATGCAAGTTC CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTCACGCTCCAGCCTCCT TCTTGGTTGTCATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA TGGGAACCAGGTCTGAAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA ${\tt CGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT}$ GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAC CTCATTTATAAAAGCTTCAAAAAAACCCA





></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624</pre>

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSD PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRK EIDEIVIELTVQVKPVTPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGG VLVVLAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267